Signomial and Polynomial Optimization via Relative Entropy and Partial Dualization

Riley Murray

Presented at MPI MiS
Leipzig, Germany

2 October 2019

Joint work with Venkat Chandrasekaran and Adam Wierman.

Definitions from convex analysis

Caltech

$X \subset \mathbb{R}^n$ is a **convex set** if it contains all of its line segments.

$f : \mathbb{R}^n \to \mathbb{R}$ is a **convex function** if

$$f(t\boldsymbol{x} + (1-t)\boldsymbol{y}) \leq tf(\boldsymbol{x}) + (1-t)f(\boldsymbol{y}).$$

for all $\boldsymbol{x}, \boldsymbol{y} \in \operatorname{dom} f$ and all $t \in [0, 1]$.

Definitions from convex analysis

Caltech

$X \subset \mathbb{R}^n$ is a **convex set** if it contains all of its line segments.

$f : \mathbb{R}^n \to \mathbb{R}$ is a **convex function** if

$$f(t\boldsymbol{x} + (1-t)\boldsymbol{y}) \leq tf(\boldsymbol{x}) + (1-t)f(\boldsymbol{y}).$$

for all $\boldsymbol{x}, \boldsymbol{y} \in \operatorname{dom} f$ and all $t \in [0,1]$.

A convex set X induces a **support function**

$$\sigma_X(\boldsymbol{\lambda}) = \sup\{\boldsymbol{\lambda}^\intercal \boldsymbol{x} \,:\, \boldsymbol{x} \text{ in } X\}.$$

Definitions from convex analysis

Caltech

$X \subset \mathbb{R}^n$ is a **convex set** if it contains all of its line segments.

$f : \mathbb{R}^n \to \mathbb{R}$ is a **convex function** if

$$f(t\boldsymbol{x} + (1-t)\boldsymbol{y}) \leq t f(\boldsymbol{x}) + (1-t) f(\boldsymbol{y}).$$

for all $\boldsymbol{x}, \boldsymbol{y} \in \operatorname{dom} f$ and all $t \in [0,1]$.

A convex set X induces a **support function**

$$\sigma_X(\boldsymbol{\lambda}) = \sup\{\boldsymbol{\lambda}^\intercal \boldsymbol{x} \,:\, \boldsymbol{x} \text{ in } X\}.$$

The **relative entropy function** continuously extends

$$D(\boldsymbol{u}, \boldsymbol{v}) = \sum_{i=1}^{m} u_i \log(u_i/v_i) \quad \text{to} \quad \mathbb{R}^m_+ \times \mathbb{R}^m_+.$$

Convex duality

Caltech

Start with a **primal** problem

$$\mathrm{Val}(\boldsymbol{c}) = \inf_{\boldsymbol{x}} \{ \boldsymbol{c}^\intercal \boldsymbol{x} \,:\, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0} \}.$$

Convex duality

Caltech

Start with a **primal** problem

$$\mathrm{Val}(\boldsymbol{c}) = \inf_{\boldsymbol{x}} \{\boldsymbol{c}^\intercal \boldsymbol{x} \,:\, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}.$$

Obtain a **dual** problem

$$\mathrm{Val}(\boldsymbol{c}) = \sup_{\boldsymbol{\mu}} \{\boldsymbol{b}^\intercal \boldsymbol{\mu} \,:\, \boldsymbol{A}^\intercal \boldsymbol{\mu} \leq \boldsymbol{c}\}.$$

Convex duality

Caltech

Start with a **primal** problem

$$\mathrm{Val}(\boldsymbol{c}) = \inf_{\boldsymbol{x}} \{\boldsymbol{c}^\intercal \boldsymbol{x} \,:\, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}.$$

Obtain a **dual** problem

$$\mathrm{Val}(\boldsymbol{c}) = \sup_{\boldsymbol{\mu}} \{\boldsymbol{b}^\intercal \boldsymbol{\mu} \,:\, \boldsymbol{A}^\intercal \boldsymbol{\mu} \leq \boldsymbol{c}\}.$$

We will encounter constraints like

$$\mathrm{Val}(\boldsymbol{c}) \geq L.$$

Convex duality

Caltech

Start with a **primal** problem

$$\mathrm{Val}(\boldsymbol{c}) = \inf_{\boldsymbol{x}}\{\boldsymbol{c}^\intercal \boldsymbol{x} \,:\, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}\}.$$

Obtain a **dual** problem

$$\mathrm{Val}(\boldsymbol{c}) = \sup_{\boldsymbol{\mu}}\{\boldsymbol{b}^\intercal \boldsymbol{\mu} \,:\, \boldsymbol{A}^\intercal \boldsymbol{\mu} \leq \boldsymbol{c}\}.$$

We will encounter constraints like

$$\mathrm{Val}(\boldsymbol{c}) \geq L.$$

Write such a constraint as: *there exists* a $\boldsymbol{\mu}$ where

$$\boldsymbol{A}^\intercal \boldsymbol{\mu} \leq \boldsymbol{c} \quad \text{and} \quad \boldsymbol{b}^\intercal \boldsymbol{\mu} \geq L.$$

Nonnegativity and Optimization

Caltech

Start with a function f and a set $X \subset \mathbb{R}^n$

$$\inf\{f(\boldsymbol{x}) : \boldsymbol{x} \text{ in } X\} = \sup\{\gamma : f(\boldsymbol{x}) \geq \gamma \text{ for all } \boldsymbol{x} \text{ in } X\}.$$

Nonnegativity and Optimization

Caltech

Start with a function f and a set $X \subset \mathbb{R}^n$

$$\inf\{f(\boldsymbol{x}) : \boldsymbol{x} \text{ in } X\} = \sup\{\gamma : f(\boldsymbol{x}) \geq \gamma \text{ for all } \boldsymbol{x} \text{ in } X\}.$$

Plan of attack for producing lower bounds:

1. Express $f - \gamma = \sum_{i=1}^{m} c_i(\gamma) \cdot \phi_i$ for some basis functions ϕ_i.

Nonnegativity and Optimization

Caltech

Start with a function f and a set $X \subset \mathbb{R}^n$

$$\inf\{f(\boldsymbol{x}) : \boldsymbol{x} \text{ in } X\} = \sup\{\gamma : f(\boldsymbol{x}) \geq \gamma \text{ for all } \boldsymbol{x} \text{ in } X\}.$$

Plan of attack for producing lower bounds:

1. Express $f - \gamma = \sum_{i=1}^m c_i(\gamma) \cdot \phi_i$ for some basis functions ϕ_i.
2. Develop an inner approximation

$$\tilde{C}(\phi, X) \subset \left\{ \boldsymbol{c} \,:\, \sum_{i=1}^m c_i \phi_i(\boldsymbol{x}) \geq 0 \text{ for all } \boldsymbol{x} \text{ in } X \right\}.$$

Nonnegativity and Optimization

Caltech

Start with a function f and a set $X \subset \mathbb{R}^n$

$$\inf\{f(\boldsymbol{x}) : \boldsymbol{x} \text{ in } X\} = \sup\{\gamma : f(\boldsymbol{x}) \geq \gamma \text{ for all } \boldsymbol{x} \text{ in } X\}.$$

Plan of attack for producing lower bounds:

1. Express $f - \gamma = \sum_{i=1}^{m} c_i(\gamma) \cdot \phi_i$ for some basis functions ϕ_i.
2. Develop an inner approximation

$$\tilde{C}(\phi, X) \subset \left\{ \boldsymbol{c} \,:\, \sum_{i=1}^{m} c_i \phi_i(\boldsymbol{x}) \geq 0 \text{ for all } \boldsymbol{x} \text{ in } X \right\}.$$

3. Find largest γ so $\boldsymbol{c}(\gamma)$ belongs to $\tilde{C}(\phi, X)$.

Our functions of interest

Caltech

polynomials

signomials

Our functions of interest

Caltech

polynomials

signomials

Parameters $\boldsymbol{\alpha}_i$ in $\mathbb{N}^n$, c_i in $\mathbb{R}$.

Using $\boldsymbol{x}^{\boldsymbol{\alpha}_i} = \prod_{j=1}^n x_j^{\alpha_{ij}}$,

$$\boldsymbol{x} \mapsto \sum_{i=1}^m c_i \boldsymbol{x}^{\boldsymbol{\alpha}_i}.$$

Our functions of interest

Caltech

polynomials

Parameters $\boldsymbol{\alpha}_i$ in $\mathbb{N}^n$, c_i in $\mathbb{R}$.

Using $\boldsymbol{x}^{\boldsymbol{\alpha}_i} = \prod_{j=1}^n x_j^{\alpha_{ij}}$,

$$\boldsymbol{x} \mapsto \sum_{i=1}^m c_i \boldsymbol{x}^{\boldsymbol{\alpha}_i}.$$

signomials

Parameters $\boldsymbol{\alpha}_i$ in $\mathbb{R}^n$, c_i in $\mathbb{R}$.

Our functions of interest

Caltech

polynomials

Parameters $\boldsymbol{\alpha}_i$ in $\mathbb{N}^n$, c_i in $\mathbb{R}$.

Using $\boldsymbol{x}^{\boldsymbol{\alpha}_i} = \prod_{j=1}^n x_j^{\alpha_{ij}}$,

$$\boldsymbol{x} \mapsto \sum_{i=1}^m c_i \boldsymbol{x}^{\boldsymbol{\alpha}_i}.$$

signomials

Parameters $\boldsymbol{\alpha}_i$ in $\mathbb{R}^n$, c_i in $\mathbb{R}$.

In “exponential form”,

$$\boldsymbol{x} \mapsto \sum_{i=1}^m c_i \exp(\boldsymbol{\alpha}_i \cdot \boldsymbol{x}).$$

Our functions of interest

Caltech

polynomials

Parameters $\boldsymbol{\alpha}_i$ in $\mathbb{N}^n$, c_i in $\mathbb{R}$.

Using $\boldsymbol{x}^{\boldsymbol{\alpha}_i} = \prod_{j=1}^n x_j^{\alpha_{ij}}$,

$$\boldsymbol{x} \mapsto \sum_{i=1}^m c_i \boldsymbol{x}^{\boldsymbol{\alpha}_i}.$$

Countable basis.

signomials

Parameters $\boldsymbol{\alpha}_i$ in $\mathbb{R}^n$, c_i in $\mathbb{R}$.

In "exponential form",

$$\boldsymbol{x} \mapsto \sum_{i=1}^m c_i \exp(\boldsymbol{\alpha}_i \cdot \boldsymbol{x}).$$

Uncountable basis.

Our functions of interest

Caltech

polynomials

Parameters $\boldsymbol{\alpha}_i$ in $\mathbb{N}^n$, c_i in $\mathbb{R}$.

Using $\boldsymbol{x}^{\boldsymbol{\alpha}_i} = \prod_{j=1}^n x_j^{\alpha_{ij}}$,

$$\boldsymbol{x} \mapsto \sum_{i=1}^m c_i \boldsymbol{x}^{\boldsymbol{\alpha}_i}.$$

Countable basis.

Complexity measured by degree $\max_i \alpha_{i1} + \cdots + \alpha_{in}$.

signomials

Parameters $\boldsymbol{\alpha}_i$ in $\mathbb{R}^n$, c_i in $\mathbb{R}$.

In "exponential form",

$$\boldsymbol{x} \mapsto \sum_{i=1}^m c_i \exp(\boldsymbol{\alpha}_i \cdot \boldsymbol{x}).$$

Uncountable basis.

Complexity measured by number of terms m.

How we think of signomials

Caltech

Signomials are often written $\boldsymbol{y} \mapsto \sum_{i=1}^{m} c_i \boldsymbol{y}^{\boldsymbol{\alpha}_i}$, with $\boldsymbol{y} \in \mathbb{R}^n_{++}$.

The *exponential form* has a powerful connection to convexity.

How we think of signomials

Caltech

Signomials are often written $\boldsymbol{y} \mapsto \sum_{i=1}^{m} c_i \boldsymbol{y}^{\boldsymbol{\alpha}_i}$, with $\boldsymbol{y} \in \mathbb{R}^n_{++}$.

The *exponential form* has a powerful connection to convexity.

One use of your existing intuition:

- Pick an "interesting" polynomial p.
- Define $f(\boldsymbol{z}) = p\left(\exp(\boldsymbol{z}) - \exp(-\boldsymbol{z})\right)$.
- f will behave similarly to p, near $\mathbf{0}$.

How we think of signomials

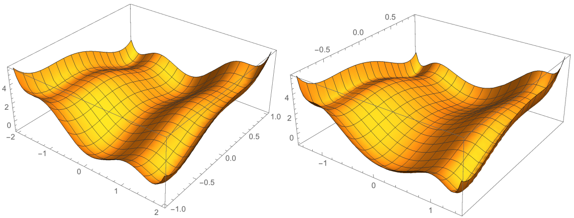
Caltech

Signomials are often written $\boldsymbol{y} \mapsto \sum_{i=1}^{m} c_i \boldsymbol{y}^{\boldsymbol{\alpha}_i}$, with $\boldsymbol{y} \in \mathbb{R}^n_{++}$.

The *exponential form* has a powerful connection to convexity.

One use of your existing intuition:

- Pick an "interesting" polynomial p.
- Define $f(\boldsymbol{z}) = p\left(\exp(\boldsymbol{z}) - \exp(-\boldsymbol{z})\right)$.
- f will behave similarly to p, near $\mathbf{0}$.



Outline for the talk

This talk is about “**SAGE certificates**.”

Outline for the talk

Caltech

This talk is about "**SAGE certificates**."

1. SAGE signomials

Definition $\rightarrow$ Representation $\rightarrow$ Example.

Outline for the talk

Caltech

This talk is about "**SAGE certificates**."

1. SAGE signomials

   Definition $\rightarrow$ Representation $\rightarrow$ Example.

2. Signomial optimization
   - Simple "SAGE relaxations."
   - Partial dualization.
   - Two examples.

Outline for the talk

Caltech

This talk is about "**SAGE certificates**."

1. SAGE signomials

   Definition $\rightarrow$ Representation $\rightarrow$ Example.

2. Signomial optimization
   - Simple "SAGE relaxations."
   - Partial dualization.
   - Two examples.

3. SAGE polynomials

   Definition $\rightarrow$ Example $\rightarrow$ Representation.

The signomial X-nonnegativity cones

Caltech

The X-nonnegativity cone for signomials over exponents $\boldsymbol{\alpha}$:

$$C_{\mathrm{NNS}}(\boldsymbol{\alpha}, X) \doteq \left\{ \boldsymbol{c} \, : \, \sum_{i=1}^{m} c_i \exp(\boldsymbol{\alpha}_i \cdot \boldsymbol{x}) \geq 0 \text{ for all } \boldsymbol{x} \text{ in } X \right\}.$$

The signomial X-nonnegativity cones

Caltech

The X-nonnegativity cone for signomials over exponents $\boldsymbol{\alpha}$:

$$C_{\mathrm{NNS}}(\boldsymbol{\alpha}, X) \doteq \left\{ \boldsymbol{c} \,:\, \sum_{i=1}^{m} c_i \exp(\boldsymbol{\alpha}_i \cdot \boldsymbol{x}) \geq 0 \text{ for all } \boldsymbol{x} \text{ in } X \right\}.$$

If

$$f(\boldsymbol{x}) = c_1 \exp(\mathbf{0} \cdot \boldsymbol{x}) + c_2 \exp(\boldsymbol{\alpha}_2 \cdot \boldsymbol{x}) + \cdots + c_m \exp(\boldsymbol{\alpha}_m \cdot \boldsymbol{x}),$$

then

$$\inf_{\boldsymbol{x} \in X} f(\boldsymbol{x}) = \sup \left\{ \gamma \,:\, \boldsymbol{c} - (\gamma, 0, \ldots, 0) \in C_{\mathrm{NNS}}(\boldsymbol{\alpha}, X) \right\}.$$

X-SAGE $\Rightarrow$ X-nonnegativity

Caltech

Definition. A signomial which is nonnegative over X and which has at most one negative coefficient is an "X-AGE function."

X-SAGE $\Rightarrow$ X-nonnegativity

Caltech

Definition. A signomial which is nonnegative over X and which has at most one negative coefficient is an "X-AGE function."

We take sums of X-AGE cones to obtain the X-**SAGE cone**

$$C_{\mathrm{SAGE}}(\boldsymbol{\alpha}, X) = \sum_{k=1}^{m} \underbrace{\{\boldsymbol{c} \ : \ \boldsymbol{c}_{\setminus k} \geq \mathbf{0} \text{ and } \boldsymbol{c} \text{ in } C_{\mathrm{NNS}}(\boldsymbol{\alpha}, X)\}}_{k^{\text{th}} \ X\text{-AGE cone}}.$$

X-SAGE $\Rightarrow$ X-nonnegativity

Caltech

Definition. A signomial which is nonnegative over X and which has at most one negative coefficient is an "X-AGE function."

We take sums of X-AGE cones to obtain the X-**SAGE cone**

$$C_{\mathrm{SAGE}}(\boldsymbol{\alpha}, X) = \sum_{k=1}^{m} \underbrace{\{\boldsymbol{c} \,:\, \boldsymbol{c}_{\setminus k} \geq \mathbf{0} \text{ and } \boldsymbol{c} \text{ in } C_{\mathrm{NNS}}(\boldsymbol{\alpha}, X)\}}_{k^{\text{th}}\ X\text{-AGE cone}}.$$

Crucial question: How to represent the AGE cones?

The convex duality behind AGE cones

Caltech

Fix $\boldsymbol{\alpha}$ in $\mathbb{R}^{m \times n}$, and $\boldsymbol{c}$ in $\mathbb{R}^m$ satisfying $\boldsymbol{c}_{\setminus k} \geq \mathbf{0}$. Convex $X \subset \mathbb{R}^n$.

The convex duality behind AGE cones

Caltech

Fix $\boldsymbol{\alpha}$ in $\mathbb{R}^{m \times n}$, and $\boldsymbol{c}$ in $\mathbb{R}^m$ satisfying $\boldsymbol{c}_{\setminus k} \geq \mathbf{0}$. Convex $X \subset \mathbb{R}^n$.

Divide out the problematic exponential, and rearrange terms:

$$\begin{aligned} \textstyle\sum_{i=1}^m c_i \exp(\boldsymbol{\alpha}_i \cdot \boldsymbol{x}) \geq 0 \;\Leftrightarrow\; & \textstyle\sum_{i=1}^m c_i \exp([\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_k] \cdot \boldsymbol{x}) \geq 0 \\ & \textstyle\sum_{i \neq k} c_i \exp([\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_k] \cdot \boldsymbol{x}) \geq -c_k. \end{aligned}$$

The convex duality behind AGE cones

Caltech

Fix $\boldsymbol{\alpha}$ in $\mathbb{R}^{m \times n}$, and $\boldsymbol{c}$ in $\mathbb{R}^m$ satisfying $\boldsymbol{c}_{\setminus k} \geq \mathbf{0}$. Convex $X \subset \mathbb{R}^n$.

Divide out the problematic exponential, and rearrange terms:

$$\sum_{i=1}^m c_i \exp(\boldsymbol{\alpha}_i \cdot \boldsymbol{x}) \geq 0 \;\Leftrightarrow\; \sum_{i=1}^m c_i \exp([\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_k] \cdot \boldsymbol{x}) \geq 0$$

$$\sum_{i \neq k} c_i \exp([\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_k] \cdot \boldsymbol{x}) \geq -c_k.$$

The nonnegativity condition

$$\inf \left\{ \sum_{i \neq k} c_i \exp([\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_k] \cdot \boldsymbol{x}) \;:\; \boldsymbol{x} \text{ in } X \right\} \geq -c_k$$

holds **if and only if**

The convex duality behind AGE cones

Caltech

Fix $\boldsymbol{\alpha}$ in $\mathbb{R}^{m \times n}$, and $\boldsymbol{c}$ in $\mathbb{R}^m$ satisfying $\boldsymbol{c}_{\setminus k} \geq \mathbf{0}$. Convex $X \subset \mathbb{R}^n$.

Divide out the problematic exponential, and rearrange terms:

$$\sum_{i=1}^m c_i \exp(\boldsymbol{\alpha}_i \cdot \boldsymbol{x}) \geq 0 \;\Leftrightarrow\; \sum_{i=1}^m c_i \exp([\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_k] \cdot \boldsymbol{x}) \geq 0$$

$$\sum_{i \neq k} c_i \exp([\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_k] \cdot \boldsymbol{x}) \geq -c_k.$$

The nonnegativity condition

$$\inf \left\{ \sum_{i \neq k} c_i \exp([\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_k] \cdot \boldsymbol{x}) \;:\; \boldsymbol{x} \text{ in } X \right\} \geq -c_k$$

holds **if and only if** there exists $\boldsymbol{\nu}$ in $\mathbb{R}^{m-1}$, $\boldsymbol{\lambda}$ in $\mathbb{R}^n$ satisfying

The convex duality behind AGE cones

Caltech

Fix $\boldsymbol{\alpha}$ in $\mathbb{R}^{m \times n}$, and $\boldsymbol{c}$ in $\mathbb{R}^m$ satisfying $\boldsymbol{c}_{\setminus k} \geq \mathbf{0}$. Convex $X \subset \mathbb{R}^n$.

Divide out the problematic exponential, and rearrange terms:

$$\sum_{i=1}^m c_i \exp(\boldsymbol{\alpha}_i \cdot \boldsymbol{x}) \geq 0 \;\Leftrightarrow\; \sum_{i=1}^m c_i \exp([\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_k] \cdot \boldsymbol{x}) \geq 0$$
$$\sum_{i \neq k} c_i \exp([\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_k] \cdot \boldsymbol{x}) \geq -c_k.$$

The nonnegativity condition

$$\inf \left\{ \sum_{i \neq k} c_i \exp([\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_k] \cdot \boldsymbol{x}) \;:\; \boldsymbol{x} \text{ in } X \right\} \geq -c_k$$

holds **if and only if** there exists $\boldsymbol{\nu}$ in $\mathbb{R}^{m-1}$, $\boldsymbol{\lambda}$ in $\mathbb{R}^n$ satisfying

$$[\mathbf{1}\boldsymbol{\alpha}_k - \boldsymbol{\alpha}_{\setminus k}]^\intercal \boldsymbol{\nu} = \boldsymbol{\lambda} \quad \text{and}$$
$$\sigma_X(\boldsymbol{\lambda}) + D(\boldsymbol{\nu}, \boldsymbol{c}_{\setminus k}) - \boldsymbol{\nu}^\intercal \mathbf{1} \leq c_k.$$

Tractability of X-SAGE cones

Caltech

There are two constraints in an AGE cone:

- $[\mathbf{1}\boldsymbol{\alpha}_k - \boldsymbol{\alpha}_{\setminus k}]^\intercal \boldsymbol{\nu} = \boldsymbol{\lambda}$, and
- $\sigma_X(\boldsymbol{\lambda}) + D(\boldsymbol{\nu}, \boldsymbol{c}_{\setminus k}) - \boldsymbol{\nu}^\intercal \mathbf{1} \leq c_k$.

Support function $\sigma_X(\boldsymbol{\lambda}) = \sup\{\boldsymbol{\lambda}^\intercal \boldsymbol{x} \ : \boldsymbol{x} \text{ in } X\}$ always convex.

Tractability of X-SAGE cones

Caltech

There are two constraints in an AGE cone:

- $[\mathbf{1}\boldsymbol{\alpha}_k - \boldsymbol{\alpha}_{\setminus k}]^\intercal \boldsymbol{\nu} = \boldsymbol{\lambda}$, and
- $\sigma_X(\boldsymbol{\lambda}) + D(\boldsymbol{\nu}, \boldsymbol{c}_{\setminus k}) - \boldsymbol{\nu}^\intercal \mathbf{1} \leq c_k$.

Support function $\sigma_X(\boldsymbol{\lambda}) = \sup\{\boldsymbol{\lambda}^\intercal \boldsymbol{x} \,:\, \boldsymbol{x} \text{ in } X\}$ always convex.

When does it have a closed form? Examples include ...

$$X = \mathbb{R}^n \quad \Rightarrow \quad \sigma_X(\boldsymbol{\lambda}) = \begin{cases} 0 & \text{if } \boldsymbol{\lambda} = \mathbf{0} \\ +\infty & \text{if } \boldsymbol{\lambda} \neq \mathbf{0} \end{cases}$$

Tractability of X-SAGE cones

Caltech

There are two constraints in an AGE cone:

- $[\mathbf{1}\boldsymbol{\alpha}_k - \boldsymbol{\alpha}_{\setminus k}]^\intercal \boldsymbol{\nu} = \boldsymbol{\lambda}$, and
- $\sigma_X(\boldsymbol{\lambda}) + D(\boldsymbol{\nu}, \boldsymbol{c}_{\setminus k}) - \boldsymbol{\nu}^\intercal \mathbf{1} \leq c_k$.

Support function $\sigma_X(\boldsymbol{\lambda}) = \sup\{\boldsymbol{\lambda}^\intercal \boldsymbol{x} \,:\, \boldsymbol{x} \text{ in } X\}$ always convex.

When does it have a closed form? Examples include ...

$$X = \{\boldsymbol{x} \,:\, \|\boldsymbol{x} - \boldsymbol{a}\| \leq r\} \quad \Rightarrow \quad \sigma_X(\boldsymbol{\lambda}) = \boldsymbol{\lambda}^\intercal \boldsymbol{a} + r\|\boldsymbol{\lambda}\|_*$$

Tractability of X-SAGE cones

Caltech

There are two constraints in an AGE cone:

- $[\mathbf{1}\boldsymbol{\alpha}_k - \boldsymbol{\alpha}_{\setminus k}]^\intercal \boldsymbol{\nu} = \boldsymbol{\lambda}$, and
- $\sigma_X(\boldsymbol{\lambda}) + D(\boldsymbol{\nu}, \boldsymbol{c}_{\setminus k}) - \boldsymbol{\nu}^\intercal \mathbf{1} \leq c_k$.

Support function $\sigma_X(\boldsymbol{\lambda}) = \sup\{\boldsymbol{\lambda}^\intercal \boldsymbol{x} \, : \boldsymbol{x}$ in $X\}$ always convex.

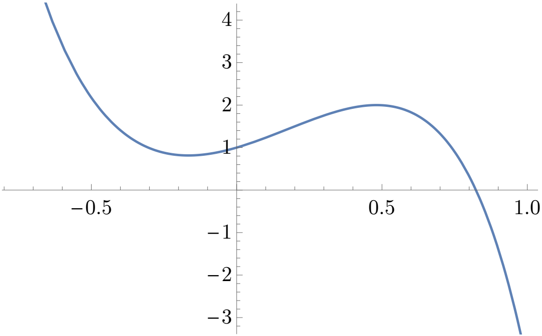
When is it *tractable*?

Whenever X is tractable.

A univariate example

Caltech

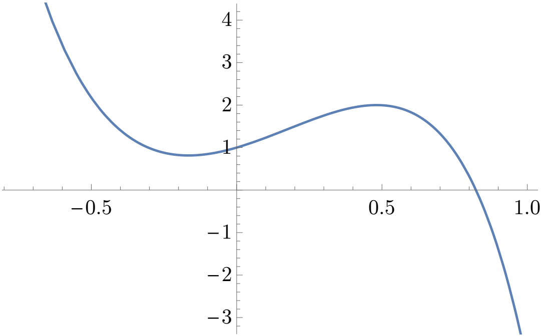
$$f(x) = e^{-3x} + e^{-2x} + 4e^{x} + e^{2x} - 4e^{-x} - 1 - e^{3x} \text{ over } x \leq 0$$



A univariate example

Caltech

$$f(x) = e^{-3x} + e^{-2x} + 4e^{x} + e^{2x} - 4e^{-x} - 1 - e^{3x} \text{ over } x \leq 0$$



$$f_1(x) = 0.88 \cdot e^{-3x} + 0.82 \cdot e^{-2x} + 2.69 \cdot e^{x} + 0.12 \cdot e^{2x} - 4 \cdot e^{-x}$$

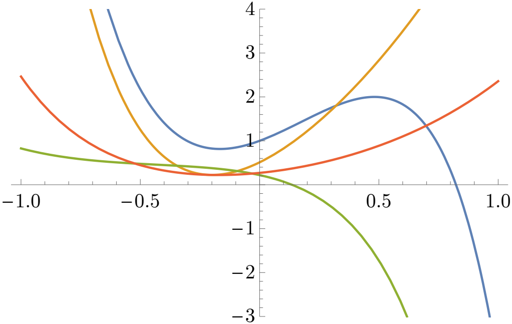
$$f_2(x) = 0.10 \cdot e^{-3x} + 0.15 \cdot e^{-2x} + 0.90 \cdot e^{x} + 0.12 \cdot e^{2x} - 1$$

$$f_3(x) = 0.02 \cdot e^{-3x} + 0.03 \cdot e^{-2x} + 0.41 \cdot e^{x} + 0.76 \cdot e^{2x} - e^{3x}$$

A univariate example

Caltech

$$f(x) = e^{-3x} + e^{-2x} + 4e^{x} + e^{2x} - 4e^{-x} - 1 - e^{3x} \text{ over } x \leq 0$$



$$f_1(x) = 0.88 \cdot e^{-3x} + 0.82 \cdot e^{-2x} + 2.69 \cdot e^{x} + 0.12 \cdot e^{2x} - 4 \cdot e^{-x}$$
$$f_2(x) = 0.10 \cdot e^{-3x} + 0.15 \cdot e^{-2x} + 0.90 \cdot e^{x} + 0.12 \cdot e^{2x} - 1$$
$$f_3(x) = 0.02 \cdot e^{-3x} + 0.03 \cdot e^{-2x} + 0.41 \cdot e^{x} + 0.76 \cdot e^{2x} - e^{3x}$$

Optimization.

Simple SAGE relaxations

Caltech

Consider $f(\boldsymbol{x}) = \sum_{i=1}^{m} c_i \exp(\boldsymbol{\alpha}_i \cdot \boldsymbol{x})$ with $\boldsymbol{\alpha}_1 = \mathbf{0}$.

$$\inf\{f(\boldsymbol{x}) \,:\, \boldsymbol{x} \text{ in } X\} = \sup\{\,\gamma \,:\, \boldsymbol{c} - \gamma \boldsymbol{e}_1 \text{ in } C_{\mathrm{NNS}}(\boldsymbol{\alpha}, X)\}$$

$$\geq \sup\{\,\gamma \,:\, \boldsymbol{c} - \gamma \boldsymbol{e}_1 \text{ in } C_{\mathrm{SAGE}}(\boldsymbol{\alpha}, X)\}$$

$$= \inf\left\{\boldsymbol{c}^\intercal \boldsymbol{v} \,:\, \begin{array}{l} \boldsymbol{v} \text{ in } C_{\mathrm{SAGE}}(\boldsymbol{\alpha}, X)^\dagger \\ \text{satisfies } \boldsymbol{v} \cdot \boldsymbol{e}_1 = 1 \end{array}\right\}$$

What about solution recovery?

Simple SAGE relaxations

Caltech

Consider $f(\boldsymbol{x}) = \sum_{i=1}^{m} c_i \exp(\boldsymbol{\alpha}_i \cdot \boldsymbol{x})$ with $\boldsymbol{\alpha}_1 = \boldsymbol{0}$.

$$\begin{aligned}\inf\{f(\boldsymbol{x}) \,:\, \boldsymbol{x} \text{ in } X\} &= \sup\{\gamma \,:\, \boldsymbol{c} - \gamma \boldsymbol{e}_1 \text{ in } C_{\mathrm{NNS}}(\boldsymbol{\alpha}, X)\} \\ &\geq \sup\{\gamma \,:\, \boldsymbol{c} - \gamma \boldsymbol{e}_1 \text{ in } C_{\mathrm{SAGE}}(\boldsymbol{\alpha}, X)\} \\ &= \inf\left\{\boldsymbol{c}^\intercal \boldsymbol{v} \,:\, \begin{array}{l} \boldsymbol{v} \text{ in } C_{\mathrm{SAGE}}(\boldsymbol{\alpha}, X)^\dagger \\ \text{satisfies } \boldsymbol{v} \cdot \boldsymbol{e}_1 = 1 \end{array}\right\}\end{aligned}$$

What about solution recovery? When X is convex, we have

$$\begin{aligned}C_{\mathrm{SAGE}}(\boldsymbol{\alpha}, X)^\dagger = \mathrm{cl}\{\boldsymbol{v} : &\text{ some } \boldsymbol{z}_1, \ldots, \boldsymbol{z}_m \text{ in } \mathbb{R}^n \text{ satisfy} \\ &v_k \log(\boldsymbol{v}/v_k) \geq [\boldsymbol{\alpha} - \boldsymbol{1}\boldsymbol{\alpha}_k]\boldsymbol{z}_k \\ &\text{and } \boldsymbol{z}_k/v_k \in X \text{ for all } k \text{ in } [m]\}.\end{aligned}$$

An example in $\mathbb{R}^3$

Minimize

$$f(\boldsymbol{x}) = 0.5\exp(x_1 - x_2) - \exp x_1 - 5\exp(-x_2)$$

over

$$\begin{aligned} X = \{\boldsymbol{x} \,:\, & \log 70 \leq x_1 \leq \log 150, \\ & \log 1.0 \leq x_2 \leq \log 30, \\ & \log 0.5 \leq x_3 \leq \log 21 \\ & \exp(x_2 - x_3) + \exp x_2 + 0.05\exp(x_1 + x_3) \leq 100\}. \end{aligned}$$

An example in $\mathbb{R}^3$

Caltech

Minimize

$$f(\boldsymbol{x}) = 0.5 \exp(x_1 - x_2) - \exp x_1 - 5 \exp(-x_2)$$

over

$$\begin{aligned} X = \{\boldsymbol{x} \, : \, & \log 70 \ \leq x_1 \leq \log 150, \\ & \log 1.0 \leq x_2 \leq \log 30, \\ & \log 0.5 \leq x_3 \leq \log 21 \\ & \exp(x_2 - x_3) + \exp x_2 + 0.05 \exp(x_1 + x_3) \leq 100\}. \end{aligned}$$

Compute $f_X^{\text{SAGE}} = -147.85713 \leq f_X^\star$,

An example in $\mathbb{R}^3$

Caltech

Minimize

$$f(\boldsymbol{x}) = 0.5\exp(x_1 - x_2) - \exp x_1 - 5\exp(-x_2)$$

over

$$\begin{aligned} X = \{\boldsymbol{x} \,:\, &\log 70 \ \leq x_1 \leq \log 150, \\ &\log 1.0 \leq x_2 \leq \log 30, \\ &\log 0.5 \leq x_3 \leq \log 21 \\ &\exp(x_2 - x_3) + \exp x_2 + 0.05\exp(x_1 + x_3) \leq 100\}. \end{aligned}$$

Compute $f_X^{\mathrm{SAGE}} = -147.85713 \leq f_X^\star$, recover feasible

$$\boldsymbol{x}^\star = (5.01063529,\ 3.40119660, -0.48450710)$$

satisfying $f(\boldsymbol{x}^\star) = -147.66666$.

An example in $\mathbb{R}^3$

Caltech

Minimize

$$f(\boldsymbol{x}) = 0.5\exp(x_1 - x_2) - \exp x_1 - 5\exp(-x_2)$$

over

$$\begin{aligned} X = \{\boldsymbol{x} \,:\, & \log 70 \ \leq x_1 \leq \log 150, \\ & \log 1.0 \leq x_2 \leq \log 30, \\ & \log 0.5 \leq x_3 \leq \log 21 \\ & \exp(x_2 - x_3) + \exp x_2 + 0.05\exp(x_1 + x_3) \leq 100\}. \end{aligned}$$

Compute $f_X^{\mathrm{SAGE}} = -147.85713 \leq f_X^\star$, recover feasible

$$\boldsymbol{x}^\star = (5.01063529,\ 3.40119660, -0.48450710)$$

satisfying $f(\boldsymbol{x}^\star) = -147.66666$. *This is actually optimal!*

Nonconvex contraints

Caltech

Q: What should we do when some constraints are nonconvex?

Nonconvex contraints

Q: What should we do when some constraints are nonconvex?

A: Combine X-SAGE certificates with Lagrangian relaxations.

Nonconvex contraints

Caltech

Q: What should we do when some constraints are nonconvex?

A: Combine X-SAGE certificates with Lagrangian relaxations.

Concretely, suppose we want to minimize f over

$$\Omega \doteq X \cap \{\boldsymbol{x} \,:\, g(\boldsymbol{x}) \leq \mathbf{0}\}$$

where X is convex, but $g_1, \ldots, g_k$ are nonconvex signomials.

Nonconvex contraints

Caltech

Q: What should we do when some constraints are nonconvex?

A: Combine X-SAGE certificates with Lagrangian relaxations.

Concretely, suppose we want to minimize f over

$$\Omega \doteq X \cap \{\boldsymbol{x} \,:\, g(\boldsymbol{x}) \leq \mathbf{0}\}$$

where X is convex, but $g_1, \ldots, g_k$ are nonconvex signomials.

Then, if $\lambda_1, \ldots, \lambda_k$ are nonnegative dual variables, we have

$$\inf_{\boldsymbol{x} \in \Omega} f(\boldsymbol{x}) \geq \sup \left\{ \gamma \,:\, f + \sum_{i=1}^{k} \lambda_i g_i - \gamma \text{ is } X\text{-SAGE} \right\}.$$

The SimPleAC aircraft design problem

Caltech

From Warren Hoburg's PhD thesis.

Problem statistics:

- 140 variables.
- 89 inequality constraints (1 nonconvex).
- 67 equality constraints (15 nonconvex).

Performance of the most basic SAGE relaxation:

- bound "cost ≥ 2957" (roughly match a known solution).
- MOSEK solves in two seconds, on a six year old laptop.
- solution recovery fails (numerical issues).

SAGE polynomials

X-nonnegative and X-AGE polynomials

Caltech

Recall our standard notation $\boldsymbol{x}^{\boldsymbol{\alpha}_i} \doteq \prod_{j=1}^{n} x_j^{\alpha_{ij}}$.

The matrix $\boldsymbol{\alpha}$ and a set $X \subset \mathbb{R}^n$ induce a nonnegativity cone

$$C_{\mathrm{NNP}}(\boldsymbol{\alpha}, X) = \{\boldsymbol{c} \,:\, c_1\boldsymbol{x}^{\boldsymbol{\alpha}_1} + \cdots + c_m\boldsymbol{x}^{\boldsymbol{\alpha}_m} \geq 0 \text{ for all } \boldsymbol{x} \text{ in } X\}.$$

X-nonnegative and X-AGE polynomials

Caltech

Recall our standard notation $\boldsymbol{x}^{\boldsymbol{\alpha}_i} \doteq \prod_{j=1}^{n} x_j^{\alpha_{ij}}$.

The matrix $\boldsymbol{\alpha}$ and a set $X \subset \mathbb{R}^n$ induce a nonnegativity cone

$$C_{\mathrm{NNP}}(\boldsymbol{\alpha}, X) = \{\boldsymbol{c} \,:\, c_1\boldsymbol{x}^{\boldsymbol{\alpha}_1} + \cdots + c_m\boldsymbol{x}^{\boldsymbol{\alpha}_m} \geq 0 \text{ for all } \boldsymbol{x} \text{ in } X\}.$$

Def. $f(\boldsymbol{x}) = c_1\boldsymbol{x}^{\boldsymbol{\alpha}_1} + \cdots + c_m\boldsymbol{x}^{\boldsymbol{\alpha}_m}$ is an X-**AGE polynomial** if

1. $\boldsymbol{c}$ belongs to $C_{\mathrm{NNP}}(\boldsymbol{\alpha}, X)$, and
2. at most one "k" has $c_k\boldsymbol{x}^{\boldsymbol{\alpha}_k} < 0$ for some $\boldsymbol{x} \in X$.

A symbolic example

Caltech

Consider $$f(\boldsymbol{x}) = n - \sum_{i=1}^{n} \prod_{j \in [n] \setminus \{i\}} x_j \quad \text{over} \quad X = \{-1, 1\}^n.$$

Let's prove that f is X-nonnegative.

A symbolic example

Caltech

Consider $$f(\boldsymbol{x}) = n - \sum_{i=1}^{n} \prod_{j \in [n] \setminus \{i\}} x_j \quad \text{over} \quad X = \{-1, 1\}^n.$$

Let's prove that f is X-nonnegative.

1. Note that $\boldsymbol{x} \in \{-1, 1\}^n$ implies $\prod_{j \neq i} x_j \leq 1$.

A symbolic example

Caltech

Consider $$f(\boldsymbol{x}) = n - \sum_{i=1}^{n} \prod_{j \in [n] \setminus \{i\}} x_j \quad \text{over} \quad X = \{-1, 1\}^n.$$

Let's prove that f is X-nonnegative.

1. Note that $\boldsymbol{x} \in \{-1, 1\}^n$ implies $\prod_{j \neq i} x_j \leq 1$.
2. Therefore $f_i(\boldsymbol{x}) = 1 - \prod_{j \neq i} x_j$ are X-AGE.

A symbolic example

Caltech

Consider $$f(\boldsymbol{x}) = n - \sum_{i=1}^{n} \prod_{j \in [n] \setminus \{i\}} x_j$$ over $X = \{-1, 1\}^n$.

Let's prove that f is X-nonnegative.

1. Note that $\boldsymbol{x} \in \{-1, 1\}^n$ implies $\prod_{j \neq i} x_j \leq 1$.
2. Therefore $f_i(\boldsymbol{x}) = 1 - \prod_{j \neq i} x_j$ are X-AGE.
3. Since $f = \sum_{i=1}^{n} f_i$, conclude f is X-SAGE.

The representation problem

Caltech

The cone of coefficients for X-**SAGE polynomials** is given by

$$C_{\mathrm{SAGE}}^{\mathrm{POLY}}(\boldsymbol{\alpha}, X) = \sum_{k=1}^{m} \left\{ \boldsymbol{c} \;\middle|\; \begin{array}{c} \boldsymbol{c} \text{ in } C_{\mathrm{NNP}}(\boldsymbol{\alpha}, X), \text{ and for} \\ i \neq k, \boldsymbol{x} \in X \Rightarrow c_i \boldsymbol{x}^{\boldsymbol{\alpha}_i} \geq 0 \end{array} \right\}.$$

The representation problem

Caltech

The cone of coefficients for X-**SAGE polynomials** is given by

$$C_{\mathrm{SAGE}}^{\mathrm{POLY}}(\boldsymbol{\alpha}, X) = \sum_{k=1}^{m} \left\{ \boldsymbol{c} \;\middle|\; \begin{array}{c} \boldsymbol{c} \text{ in } C_{\mathrm{NNP}}(\boldsymbol{\alpha}, X), \text{ and for} \\ i \neq k, \boldsymbol{x} \in X \Rightarrow c_i \boldsymbol{x}^{\boldsymbol{\alpha}_i} \geq 0 \end{array} \right\}.$$

- This definition cares about the behavior of monomials.

The representation problem

Caltech

The cone of coefficients for X-**SAGE polynomials** is given by

$$C_{\mathrm{SAGE}}^{\mathrm{POLY}}(\boldsymbol{\alpha}, X) = \sum_{k=1}^{m} \left\{ \boldsymbol{c} \;\middle|\; \begin{array}{c} \boldsymbol{c} \text{ in } C_{\mathrm{NNP}}(\boldsymbol{\alpha}, X), \text{ and for} \\ i \neq k, \boldsymbol{x} \in X \Rightarrow c_i \boldsymbol{x}^{\boldsymbol{\alpha}_i} \geq 0 \end{array} \right\}.$$

- This definition cares about the behavior of monomials.
- SAGE signomials put monomials front and center.

The representation problem

Caltech

The cone of coefficients for X-**SAGE polynomials** is given by

$$C_{\text{SAGE}}^{\text{POLY}}(\boldsymbol{\alpha}, X) = \sum_{k=1}^{m} \left\{ \boldsymbol{c} \;\middle|\; \begin{array}{c} \boldsymbol{c} \text{ in } C_{\text{NNP}}(\boldsymbol{\alpha}, X), \text{ and for} \\ i \neq k, \boldsymbol{x} \in X \Rightarrow c_i \boldsymbol{x}^{\boldsymbol{\alpha}_i} \geq 0 \end{array} \right\}.$$

- This definition cares about the behavior of monomials.
- SAGE signomials put monomials front and center.
- Monomials "$\boldsymbol{x}^{\boldsymbol{\alpha}_i}$" are quite different from "$\exp(\boldsymbol{\alpha}_i \cdot \boldsymbol{x})$".

Representation: the single-orthant case

Caltech

Consider $X \subset \mathbb{R}^n_+$, and define $X_{++} \doteq X \cap \mathbb{R}^n_{++}$.

$$C^{\mathrm{POLY}}_{\mathrm{SAGE}}(\boldsymbol{\alpha}, X) = \sum_{k=1}^{m} \left\{ \boldsymbol{c} \;\middle|\; \begin{array}{l} \boldsymbol{c} \text{ in } C_{\mathrm{NNP}}(\boldsymbol{\alpha}, X), \text{ and for} \\ i \neq k, \boldsymbol{x} \in X \Rightarrow c_i \boldsymbol{x}^{\boldsymbol{\alpha}_i} \geq 0 \end{array} \right\}$$

Representation: the single-orthant case

Caltech

Consider $X \subset \mathbb{R}^n_+$, and define $X_{++} \doteq X \cap \mathbb{R}^n_{++}$.

$$C^{\text{POLY}}_{\text{SAGE}}(\boldsymbol{\alpha}, X) = \sum_{k=1}^{m} \left\{ \boldsymbol{c} \;\middle|\; \begin{array}{c} \boldsymbol{c} \text{ in } C_{\text{NNP}}(\boldsymbol{\alpha}, X), \text{ and for} \\ i \neq k, \boldsymbol{x} \in X \Rightarrow c_i \boldsymbol{x}^{\boldsymbol{\alpha}_i} \geq 0 \end{array} \right\}$$

Fact 1. If $X = \operatorname{cl} X_{++}$, then $C_{\text{NNP}}(\boldsymbol{\alpha}, X) = C_{\text{NNS}}(\boldsymbol{\alpha}, \log X_{++})$.

Representation: the single-orthant case

Caltech

Consider $X \subset \mathbb{R}^n_+$, and define $X_{++} \doteq X \cap \mathbb{R}^n_{++}$.

$$C^{\mathrm{POLY}}_{\mathrm{SAGE}}(\boldsymbol{\alpha}, X) = \sum_{k=1}^{m} \left\{ \boldsymbol{c} \;\middle|\; \begin{array}{l} \boldsymbol{c} \text{ in } C_{\mathrm{NNP}}(\boldsymbol{\alpha}, X), \text{ and for} \\ i \neq k, \boldsymbol{x} \in X \Rightarrow c_i \boldsymbol{x}^{\boldsymbol{\alpha}_i} \geq 0 \end{array} \right\}$$

Fact 1. If $X = \operatorname{cl} X_{++}$, then $C_{\mathrm{NNP}}(\boldsymbol{\alpha}, X) = C_{\mathrm{NNS}}(\boldsymbol{\alpha}, \log X_{++})$.

Fact 2. Since $X \subset \mathbb{R}^n_+$, "$c_i \boldsymbol{x}^{\boldsymbol{\alpha}_i} \geq 0 \; \forall \; \boldsymbol{x} \in X$" reduces to $c_i \geq 0$.

Representation: the single-orthant case

Caltech

Consider $X \subset \mathbb{R}^n_+$, and define $X_{++} \doteq X \cap \mathbb{R}^n_{++}$.

$$C^{\mathrm{POLY}}_{\mathrm{SAGE}}(\boldsymbol{\alpha}, X) = \sum_{k=1}^{m} \left\{ \boldsymbol{c} \;\middle|\; \begin{array}{l} \boldsymbol{c} \text{ in } C_{\mathrm{NNP}}(\boldsymbol{\alpha}, X), \text{ and for} \\ i \neq k, \boldsymbol{x} \in X \Rightarrow c_i \boldsymbol{x}^{\boldsymbol{\alpha}_i} \geq 0 \end{array} \right\}$$

Fact 1. If $X = \mathrm{cl}\, X_{++}$, then $C_{\mathrm{NNP}}(\boldsymbol{\alpha}, X) = C_{\mathrm{NNS}}(\boldsymbol{\alpha}, \log X_{++})$.

Fact 2. Since $X \subset \mathbb{R}^n_+$, "$c_i \boldsymbol{x}^{\boldsymbol{\alpha}_i} \geq 0 \;\forall\; \boldsymbol{x} \in X$" reduces to $c_i \geq 0$.

$$\sum_{k=1}^{m} \left\{ \boldsymbol{c} \;\middle|\; \begin{array}{l} \boldsymbol{c} \text{ in } C_{\mathrm{NNS}}(\boldsymbol{\alpha}, \log X_{++}), \\ \text{and } c_i \geq 0 \text{ for all } i \neq k \end{array} \right\} = C_{\mathrm{SAGE}}(\boldsymbol{\alpha}, \log X_{++})$$

Representation: sign-symmetry

Caltech

Consider $X \subset \mathbb{R}^n$ which satisfies

1. invariance under reflection about $\{\boldsymbol{x} : x_i = 0\}$,
2. and $X \cap \mathbb{R}^n_+ = \mathrm{cl}\, X_{++}$.

For example, $X = \{-1, 1\}^n$.

Representation: sign-symmetry

Caltech

Consider $X \subset \mathbb{R}^n$ which satisfies

1. invariance under reflection about $\{\boldsymbol{x} : x_i = 0\}$,
2. and $X \cap \mathbb{R}^n_+ = \mathrm{cl}\, X_{++}$.

For example, $X = \{-1, 1\}^n$.

It can subsequently be shown that

$$C^{\mathrm{POLY}}_{\mathrm{SAGE}}(\boldsymbol{\alpha}, X) = \{\boldsymbol{c} : \text{some } \hat{\boldsymbol{c}} \in C_{\mathrm{SAGE}}(\boldsymbol{\alpha}, \log X_{++}) \text{ satisfies}$$
$$\hat{c}_i = c_i \text{ whenever } \boldsymbol{\alpha}_i \text{ is in } 2\mathbb{N}^n, \text{ and}$$
$$\hat{c}_i \leq -|c_i| \text{ whenever } \boldsymbol{\alpha}_i \text{ is not in } 2\mathbb{N}^n\}.$$

Representation: sign-symmetry

Caltech

Consider $X \subset \mathbb{R}^n$ which satisfies

1. invariance under reflection about $\{\boldsymbol{x} : x_i = 0\}$,
2. and $X \cap \mathbb{R}^n_+ = \operatorname{cl} X_{++}$.

For example, $X = \{-1, 1\}^n$.

It can subsequently be shown that

$$C^{\mathrm{POLY}}_{\mathrm{SAGE}}(\boldsymbol{\alpha}, X) = \{\boldsymbol{c} : \text{some } \hat{\boldsymbol{c}} \in C_{\mathrm{SAGE}}(\boldsymbol{\alpha}, \log X_{++}) \text{ satisfies}$$
$$\hat{c}_i = c_i \text{ whenever } \boldsymbol{\alpha}_i \text{ is in } 2\mathbb{N}^n, \text{ and}$$
$$\hat{c}_i \leq -|c_i| \text{ whenever } \boldsymbol{\alpha}_i \text{ is not in } 2\mathbb{N}^n\}.$$

If $\boldsymbol{\alpha}_i$ isn't even, then some $\boldsymbol{x}_1, \boldsymbol{x}_2 \in X$ satisfy $\boldsymbol{x}_1^{\boldsymbol{\alpha}_i} < 0 < \boldsymbol{x}_2^{\boldsymbol{\alpha}_i}$.

Log-log convexity

Caltech

For what $X \subset \mathbb{R}^n_{++}$ is $\log X$ convex?

Log-log convexity

Caltech

For what $X \subset \mathbb{R}^n_{++}$ is $\log X$ convex?

A function g is **log-log convex** on $D \subset \mathbb{R}^n_{++}$ when

- $\log D$ is a convex set, and
- $\log(\boldsymbol{x}) \mapsto \log g(\boldsymbol{x})$ is convex function.

Such functions are sometimes called *geometrically convex*.

Log-log convexity

Caltech

For what $X \subset \mathbb{R}^n_{++}$ is $\log X$ convex?

A function g is **log-log convex** on $D \subset \mathbb{R}^n_{++}$ when

- $\log D$ is a convex set, and
- $\log(\boldsymbol{x}) \mapsto \log g(\boldsymbol{x})$ is convex function.

Such functions are sometimes called *geometrically convex*.

Studied by Montel (1928) and Niculescu (2000), among others.

If $g_1, \ldots, g_k$ are log-log convex on a box $B \subset \mathbb{R}^n_{++}$, then

$$\log\{(\boldsymbol{x}, \boldsymbol{t}) \,:\, \boldsymbol{x} \in B,\ \boldsymbol{t} \in \mathbb{R}^k_{++},\ g(\boldsymbol{x}) \leq \boldsymbol{t}\} \subset \mathbb{R}^{n+k} \quad \text{is convex.}$$

Log-log convexity: examples

Caltech

With domains $D = \mathbb{R}^n_{++}$:

$$g(\boldsymbol{x}) = \max\{x_1, \ldots, x_n\}$$

Log-log convexity: examples

Caltech

With domains $D = \mathbb{R}^n_{++}$:

$$g(\boldsymbol{x}) = \max\{x_1, \ldots, x_n\}$$

$$g(\boldsymbol{x}) = x_1^{a_1} \cdots x_n^{a_n}$$

Log-log convexity: examples

Caltech

With domains $D = \mathbb{R}^n_{++}$:

$$g(\boldsymbol{x}) = \max\{x_1, \ldots, x_n\}$$

$$g(\boldsymbol{x}) = x_1^{a_1} \cdots x_n^{a_n}$$

$$g(x) = \left(\int_x^\infty e^{-t^2} \mathrm{dt}\right)^{-1}$$

Log-log convexity: examples

Caltech

With domains $D = \mathbb{R}^n_{++}$:

$$g(\boldsymbol{x}) = \max\{x_1, \ldots, x_n\}$$

$$g(\boldsymbol{x}) = x_1^{a_1} \cdots x_n^{a_n}$$

$$g(x) = \left(\int_x^\infty e^{-t^2} \mathrm{dt} \right)^{-1}$$

With more restricted domains:

$$x \mapsto (-x \log x)^{-1} \qquad D = (0, 1)$$

Log-log convexity: examples

Caltech

With domains $D = \mathbb{R}^n_{++}$:

$$g(\boldsymbol{x}) = \max\{x_1, \ldots, x_n\}$$

$$g(\boldsymbol{x}) = x_1^{a_1} \cdots x_n^{a_n}$$

$$g(x) = \left(\int_x^\infty e^{-t^2} \mathrm{dt} \right)^{-1}$$

With more restricted domains:

$$x \mapsto (-x \log x)^{-1} \qquad D = (0,1)$$

$$\boldsymbol{X} \mapsto (\boldsymbol{I} - \boldsymbol{X})^{-1}$$

$$D = \{\boldsymbol{X} \in \mathbb{R}^{n \times n}_{++} \, : \, \rho(\boldsymbol{X}) < 1\}$$

Log-log convexity: examples

Caltech

With domains $D = \mathbb{R}^n_{++}$:

$$g(\boldsymbol{x}) = \max\{x_1, \ldots, x_n\}$$

$$g(\boldsymbol{x}) = x_1^{a_1} \cdots x_n^{a_n}$$

$$g(x) = \left(\int_x^\infty e^{-t^2} \mathrm{dt} \right)^{-1}$$

With more restricted domains:

$$x \mapsto (-x \log x)^{-1} \qquad D = (0, 1)$$

$$\boldsymbol{X} \mapsto (\boldsymbol{I} - \boldsymbol{X})^{-1}$$

$$D = \{\boldsymbol{X} \in \mathbb{R}^{n \times n}_{++} \, : \, \rho(\boldsymbol{X}) < 1\}$$

$$x \mapsto (\log x)^{-1} \qquad D = (1, \infty)$$

Log-log convexity: examples

Caltech

With domains $D = \mathbb{R}^n_{++}$:

$$g(\boldsymbol{x}) = \max\{x_1, \ldots, x_n\}$$

$$g(\boldsymbol{x}) = x_1^{a_1} \cdots x_n^{a_n}$$

$$g(x) = \left(\int_x^\infty e^{-t^2} \mathrm{dt}\right)^{-1}$$

With more restricted domains:

$$x \mapsto (-x \log x)^{-1} \qquad D = (0,1)$$

$$\boldsymbol{X} \mapsto (\boldsymbol{I} - \boldsymbol{X})^{-1}$$

$$D = \{\boldsymbol{X} \in \mathbb{R}^{n \times n}_{++} \,:\, \rho(\boldsymbol{X}) < 1\}$$

$$x \mapsto (\log x)^{-1} \qquad D = (1, \infty)$$

Some tractable constraints for X-SAGE polynomials:

Log-log convexity: examples

Caltech

With domains $D = \mathbb{R}^n_{++}$:

$$g(\boldsymbol{x}) = \max\{x_1, \ldots, x_n\}$$

$$g(\boldsymbol{x}) = x_1^{a_1} \cdots x_n^{a_n}$$

$$g(x) = \left(\int_x^\infty e^{-t^2} \mathrm{dt}\right)^{-1}$$

With more restricted domains:

$$x \mapsto (-x \log x)^{-1} \qquad D = (0, 1)$$

$$\boldsymbol{X} \mapsto (\boldsymbol{I} - \boldsymbol{X})^{-1}$$

$$D = \{\boldsymbol{X} \in \mathbb{R}^{n \times n}_{++} \,:\, \rho(\boldsymbol{X}) < 1\}$$

$$x \mapsto (\log x)^{-1} \qquad D = (1, \infty)$$

Some tractable constraints for X-SAGE polynomials:

$$\|\boldsymbol{x}\|_p \leq a \qquad x_j^2 = a \qquad a \leq \mathbb{P}\{\mathcal{N}(0, \sigma) \geq |x|\}$$

where $a > 0$.

Thank you!

Handling conditional nonnegativity

Caltech

Typically, one reduces "X-nonnegativity" to the case $X = \mathbb{R}^n$.

The standard recipe

1. Adopt a representation $X = \{\boldsymbol{x} : g(\boldsymbol{x}) \geq \mathbf{0}\}$.
2. Find an identity

$$f = \mathcal{L} + \sum_i \lambda_i g_i$$

where $\mathcal{L}$ and λ_i are known to be nonnegative on $\mathbb{R}^n$.

E.g., the *positivstellensatz* of Putinar, Stengle, or Schmudgen.

Use the sageopt python package.

Caltech

- Python 3.5 or higher (recommend $\geq$ 3.6).
- "`pip install sageopt`"
- Signomial and polynomial optimization.
- Require open-source convex solver, ECOS.
- Recommend commercial solver, MOSEK.